



ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Plant Biology

Patterns of genomic changes with crop domestication and breeding

Junpeng Shi and Jinsheng Lai

Crop domestication and further breeding improvement have long been important areas of genetics and genomics studies. With the rapid advancing of next-generation sequencing (NGS) technologies, the amount of population genomics data has surged rapidly. Analyses of the mega genomics data have started to uncover a previously unknown pattern of genome-wide changes with crop domestication and breeding. Selection during domestication and breeding drastically reshaped crop genomes, which have ended up with regions of greatly reduced genetic diversity and apparent enrichment of potentially beneficial alleles located in both genic and non-genic regions. Increasing evidences suggest that epigenetic modifications also played an important role during domestication and breeding.

Addresses

State Key Laboratory of Agrobiotechnology and National Maize Improvement Center, Department of Plant Genetics and Breeding, China Agricultural University, Beijing 100193, China

Corresponding author: Lai, Jinsheng (jlai@cau.edu.cn)

Current Opinion in Plant Biology 2015, **24**:47–53

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Insuk Lee** and **Todd C Mockler**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 2nd February 2015

<http://dx.doi.org/10.1016/j.pbi.2015.01.008>

1369-5266/© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Modern crop varieties contain a number of superior agronomic traits to meet human needs and to adapt to local agronomic environments. These varieties are the products of extensive scientific breeding from landraces, which are domesticated for more than ten thousand years. Both breeding and domestication processes have been the subject of extensive genetic and genomic research. Recently, a number of insightful reviews have summarized the studies on molecular genetics changes during crop domestication and breeding [1–4].

Rapid advancements in next-generation sequencing (NGS) technology have provided a unique opportunity

for population genomics of crop domestication and breeding, because genome-wide sequencing information of large numbers of wild relatives and modern cultivars of many crops are relatively easily available. Here, we summarize the latest advances in the availability of genomic information for crop domestication and breeding improvement, the emerging methods to analyze the population genomics data and the patterns in genomics and epigenomics changes that occur with crop domestication and breeding.

Rapidly increasing population genomics data provide unprecedented opportunity for study of crop domestication and breeding

Traditionally, studies on crop domestication and breeding have been addressed using relatively small numbers of specific traits or by analyzing sequencing data of targeted regions [5]. However, the advent of NGS technologies has dramatically reduced the cost of sequencing, and so the number of crops with their entire genomes nearly completely sequenced and re-sequencing data of large numbers of individuals has increased very rapidly over the last several years. As a summary, the list of crops that have a completely sequenced genome and at the same time with reasonable data of population re-sequencing of either domesticated lines or wild relatives are shown in [Supplementary Table 1](#). Except for the major staple crops such as rice, maize and sorghum that had their reference genomes sequenced using traditional Sanger sequencing, the others were primarily sequenced using NGS. Additionally, there are larger number of crops that had their first genome reported only very recently. It is highly likely that many of these crops will have population re-sequencing efforts underway.

Coupled with the availability of large amounts of genomics data for many crops, the methods of analyzing these data have also been rapidly developed. One of the most general trends during crop domestication is a dramatically reduced genetic diversity, known as a genetic bottleneck [6,7]. Because the reduction is uneven along chromosomes, with putative selected genes experiencing more severe bottlenecks than unselected ones, such distinct genetic characteristics can be used to identify so-called selective sweeps. For a given breeding population without ancestor information, the extremely low genetic diversity (π) or Tajima's D can be used to scan the selective regions [8,9]. The composite likelihood ratio (CLR) approach has been shown to be very useful in identifying selective sweeps, which accurately predict the location

and selection coefficient of each selective sweep by taking into account complex demographics and varying mutation and recombination rates [10]. The extended haplotype homozygosity (EHH) method, which detects long-range haplotypes with unusually high EHH, can be used to find recent positive selection [11].

When the population data of both wild ancestors and modern domesticated lines are available, the selective sweeps can be identified by comparing distinct genetic characteristics between two populations. A straightforward method is to scan the genome for regions with significant reduction ($\pi_{\text{wild}}/\pi_{\text{cultivar}}$) of genetic diversity [7,12^{••},13,14]. This method surpasses the older method that simply scanned low genetic diversity regions in a single breeding population by excluding the false positive regions where there is a relatively low genetic diversity in ancestors. Population differentiation statistics, such as *F_{st}* which measures variation of Single Nucleotide Polymorphism (SNP) allele frequency between two populations, can also be used to identify differentiation selection [7,15]. Additionally, a cross-population composite likelihood ratio (XP-CLR) approach was developed, which jointly calculates

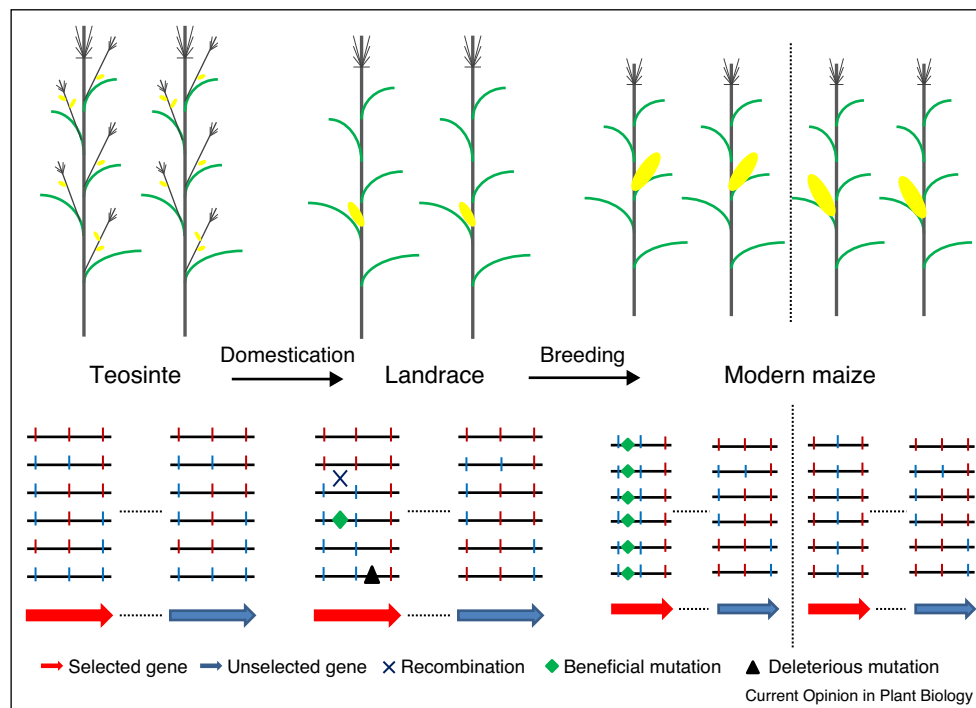
multiple locus allele frequency differentiation to identify selective sweeps between two groups [16].

The methods above are all bottom-up approaches that start with identifying genes with signatures of selection or adaptation by population genomic scanning [17]. However, these hypothetical-based strategies can unavoidably introduce many false-positives. Since most of the traits related to domestication and breeding are believed to be quantitative complex traits, the large number of selective sweeps identified from genomic scanning need to be validated using the results from traditional Quantitative Trait Locus (QTL) mapping.

Genomic changes with crop domestication and breeding

A general observation in genomics during crop domestication is the genetic bottleneck. After domestication, only favorable haplotypes are retained around selected genes (Figure 1), which creates a valley with extremely low genetic diversity. When estimated from the entire genome, π was reduced in rice from 0.003 of *Oryza rufipogon* to 0.0024 of *Oryza sativa* [12^{••}], 0.0059–0.0048 in maize

Figure 1



Genomic changes with crop domestication and breeding. *Upper plot*: morphological changes during the domestication and breeding of maize. In Teosintes, the main stalk contained multiple branches each tipped a tassel and several small ears. After domestication, maize landraces retained only one primary branch with a moderate size ear along the stalk. Modern breeding improvement process generated maize cultivars with much improved ear. *Lower plot*: genomic changes during domestication and improvement. In Teosintes, there are multiple haplotypes in both selected genes and unselected genes. After domestication, unselected genes experienced genetic bottleneck with haplotype number reduced. While in selected genes, a more severe bottleneck leave fewer haplotypes. In landrace, a beneficial mutation (green diamond) and a deleterious mutation (black triangle) arose in selected genes, along with a recombination which generated a new haplotype (red–blue–red). Breeding selection selected the newly generated haplotype (red–blue–red) and haplotype with beneficial mutation (blue–green–blue–red), together with the deleterious mutation eliminated by purifying selection. No selection happened in unselected genes during breeding.

[18^{••}], 0.0048–0.0015 in cucumber [13] and 0.0032–0.0012 in tomato [14]. The levels of genetic diversity reduction from domestication typically range from less than onefold in rice and maize, to around threefold in cucumber and tomato. However, there are regions, such as 23.8–24.9 Mb in chromosome 8 in rice, which can have more severe bottlenecks of a ~15-fold genetic diversity reduction [12^{••}]. The extent of reduction in genetic diversity during domestication is affected by the propagation system of crops. In general, much more severe genetic bottlenecks can be seen in reproductively propagated crops as comparing to that in vegetatively propagated crops [19]. For example, grape is vegetatively propagated, and was reported to have maintained a high level of genetic diversity and a rapid rate of Linkage Disequilibrium (LD) decay among domesticated individuals [20]. Genomic regions associated with reproductive barriers can have unusual levels of genetic diversity reduction. As such, a 2.2-Mb region (3.4–5.6 Mb) on chromosome 3 in watermelon was identified with a very high level of genetic diversity in wild relatives but with nearly no genetic diversity among modern cultivars [21]. A similar case was also reported in rice, suggesting that crop domestication could be responsible for the rapid evolution of reproductive barriers [22].

Despite the reduction of genetic diversity, there are also a number of cases demonstrating that population expansion, which introduced additional alleles into a defined population, can happen during the domestication and breeding processes (Figure 1). An analysis in soybean showed that low-frequency alleles were less abundant in wild accessions compared to modern cultivars [23[•]]. It was also shown that the fractions of rare alleles increased following the generations of pedigree breeding in maize, suggesting that the more advanced maize lines had accumulated higher percentages of agronomically beneficial rare alleles through recombination and selection during the breeding process [24^{••}]. There are also examples showing that functionally beneficial alleles used to be rare alleles in the ancestral population and have become common in advanced breeding varieties, such as *sh4* [25], *tga1* [26] and *fw2.2* [27]. Introduction of new alleles during the crop domestication and breeding processes can be achieved not only through the pyramiding of pre-existing alleles in the wild ancestors, but also through the keeping of newly generated alleles during the domestication and breeding processes. It has been demonstrated that genetic changes (SNPs and indels) happened rapidly with breeding [24^{••}]. There is also a report showing that a newly generated functional allele for a genotype of determinate growth habit in soybean (*GmTff1*) was artificially selected in the early stage of landrace radiation after domestication [28].

The underlying molecular mechanisms that shape crop domestication and breeding can be through either the

change of expressional levels of genes or the change of protein sequences. Although it was hypothesized that the change in gene expression might have a bigger role, the exact proportions that gene expressional changes have represented in domestication or breeding have not been extensively tested in many crops. Map-based cloning of a flowering-time QTL *vgt1* in maize showed that it is likely a putative cis-regulatory element, as it was delimited to a non-coding region upstream of the *Ap2*-like transcription factor [29]. A transposable element insertion upstream of *tb1* was reported to have interrupted an enhancer function, which at least partially explained the increased apical dominance in maize [30,31]. Results of genome-wide studies are also consistent with the notion that non-coding regulatory regions may have played important roles in crop domestication and breeding [18^{••},24^{••}]. Comparing the re-sequencing data between teosinte and modern maize, about ~6 and ~11% of the selection regions do not have coding potential, suggested that they played regulatory roles in maize domestication and breeding [18^{••}]. Similarly, a study on elite maize inbred lines suggested that many of the putative selection regions indicated by CLR analysis were located in non-genic regions [24^{••}]. A systematic tabulation of the genome wide association studies (GWAS)-associated SNPs in maize nested association mapping (NAM) population concluded that about half of the trait-associated SNPs were located in the upstream promoter regions of genes, indicating again the importance of regulatory regions in crop domestication and breeding [32[•]]. The contribution of non-genic regions in crop domestication and breeding may have been underestimated due to the low resolution of the methods used and the tight linkage between the regulatory elements and genes. Results of a number of GWAS analyses in humans seem to indicate that the majority (~93%) of trait-associated SNPs are located in non-coding regions [33].

Analysis of genomic data for wild ancestors and their modern cultivars can also reconstruct domestication or improvement events of a given crop. The domestication process is complex, and the majority of crops experienced a single domestication event, such as rice (*O. sativa* [12^{••}] and *Oryza glaberrima* [34[•]]) and maize [35]. However, some crops like common bean [36] and sorghum [37] experienced two independent domestication events. Interestingly, although both *O. sativa* and *O. glaberrima* each experienced only one geographically isolated domestication event, it was observed that for 19 domestication genes in *O. sativa*, 16 were also identified as orthologous among domestication genes in *O. glaberrima*, clearly demonstrating a convergent yet independent selection of a common gene set in two geographically distinct domestication processes [34[•]]. Two independent domestication events within a species can also have dramatically different domestication targets such that only 10% of the putative domestication sequences are shared between two domes-

tication processes in common bean [36]. A re-sequencing analysis of ~1500 *O. rufipogon* and *O. sativa* accessions resolved a long controversy about the domestication history of *O. sativa*. The final conclusion is that *O. sativa* experienced only one domestication event that *japonica* was first domesticated from *O. rufipogon* and then crossed with local wild accessions to generate *indica* [12**].

Epigenomic changes with crop domestication and breeding

For nearly all crops tested, a large proportion of their phenotypic variation cannot be fully explained by QTLs from linkage mapping or GWAS [38]. This so-called missing heritability is at least in part due to epigenetic mechanisms. Understanding the contribution of epigenetic modification in crop domestication and breeding will be crucial for further improvement of crops [39]. Heritable epigenetic changes, including DNA methylation and histone modification can modulate the genome accessibility for transcriptional machineries, therefore contribute to phenotypic variation [39,40]. In plants, DNA methylation occurs in cytosine bases of all three sequence contexts: CG, CHG and CHH [41]. Like genetic mutations, DNA methylation mutations can arise spontaneously, which create heritable epialleles with a rate of 10^{-4} to 10^{-5} methylation polymorphisms per CG site per year [42], which is much higher than the genetic mutation rate (10^{-8} to 10^{-9}) in both *Arabidopsis* [43] and maize [24**]. Such a relative high epimutation rate generates abundant epialleles with high stability. For example, the epialleles of *Fie-1* in rice can bring a dwarf stature and various floral defects without nucleotide changes but with hypomethylation, reduced H3K9me2 and increased H3K4me3 in its 5'-region [44]. Also, an epiallele of *Cnr* locating in the promoter of a SBP-factor gene controls fruit ripening in tomato [45]. Whole genome profiling of DNA methylation in wild *Arabidopsis* accessions and maize inbred lines revealed widespread existence of SMPs (single methylation polymorphisms) and DMRs (differentially methylated regions) in natural populations [42,46,47*,48*]. An important question is how many of these DNA methylation variants are controlled by genetic variation, and how many of them are purely epigenetic factors that arise spontaneously? QTL mapping for methylation variations in soybean recombinant inbred lines indicated that there are several DMRs not co-segregating with genetic variation [49], suggesting that some DNA methylation polymorphisms are independent of genetic determinants.

Several recent studies suggest that DNA methylation can also have a role in controlling heterosis. DNA methylome analysis in rice hybrids demonstrated that 0.8% of the cytosines had methylation changes between either of the parents and that in the hybrid [50]. Similarly, two studies in *Arabidopsis* revealed that DNA methylation level had increased globally in hybrids [51,52]. These DNA methylation

epimutation sites are usually covered by small RNAs, indicating a potential role of RNA-directed DNA methylation (RdDM) pathways in modulating the DNA methylation remodeling in hybrids [51]. Similar results have also been seen from the analysis of methylomes and transcriptomes of four elite Chinese maize inbred lines with known breeding pedigree information (Shaojun Xie *et al.*, unpublished).

Similar to genetic dissection of QTLs, two strategies, linkage mapping through isogenic lines [53**] and EWAS [54*] (epigenome wide association studies) analysis, can be used to find agronomically important epi-QTLs. Research on epi-QTL linkage mapping in *Arabidopsis* using isogenic lines demonstrated that six DMRs served as epi-QTLs, which accounts for a substantial (60–90%) proportion of heritability for two complex traits: flowering time and primary root length [53**]. Until now, only a few EWAS analyses have been performed in humans on traits such as smoking [55], aging related phenotypes [56], body-mass index [57] and Alzheimer's disease [58]. EWAS typically requires ~30× coverage of each sample to accurately profile DNA methylation level by whole-genome MethylC-seq, which limits large-scale population epigenetic studies in crops such as maize with relatively large genome size [59] due to sequencing costs. With sequencing costs continually declining, it is very likely that such an association analysis can be carried out for agronomic traits in major crops in the future.

Traits and underlying genes subject to selection

During domestication, conscious and unconscious selection occurred in a variety of traits such as seed shattering, plant architecture and inflorescence-related traits. Several important genes controlling these traits have been cloned using classical map-based cloning approaches, for example *PROG1* [60,61], *GIF1* [62], *tb1* [63], *tga1* [26] and *sh1* [64]. Apparently, traits controlled by a small number of large-effect QTLs are more easily domesticated. For example, the domestication of tomato from small berries to large fruit size was relatively rapid, with a major QTL *fw2.2* [27] accounting for ~30% of the phenotypic changes.

Selections during breeding are more specific on traits that are valuable to humans, although many overlap with those of domestication. Grain size and tiller number/angle are highly selected during rice breeding. As such, large numbers of QTLs related to these traits have been well characterized: *GS3* [65], *GS5* [66], *GW2* [67], *GW5* [68] and *DEP1* [69] for grain size; and *Ghd7* [70], *MOC1* [71] and *IPA1* [72] for plant architecture. It is not surprising that many of the domestication and breeding QTLs for any specific traits are functionally closely related or even biochemically interact with each other. Our future work will not only be the understanding of a particular QTL,

but also hopefully the characterization of the entire network or pathway of the important domestication or breeding traits.

Conclusions and future perspectives

The advent of genomics in recent years has provided us an unprecedented opportunity for studies of crop domestication and breeding. With large amount of genome-wide population genomics data become available, large number of selective sweeps are being identified in a high throughput manner. Genomic variations in both genic and non-genic regions were shown to be correlated with crop domestication and breeding improvement, with many underlying large effect domestication genes and breeding genes cloned. In the near future, selective sweeps will be identified with higher resolution and fewer error rates due to the availability of even larger amount of genomics data. Comparison of the potential domestication and breeding regions of closely related crop species will help to identify a number of common loci that are of special agronomic importance. Biological functions of genes locating in many of these selective regions will be understood, although they may just account for relatively small effect of domestication and breeding improvement. Of particular, more examples of non-coding functions responsible for domestication and breeding will be presented in the future. In addition, the understanding of the contribution and mechanism of epigenetic variations on crop domestication and breeding will be much improved.

Acknowledgements

We thank Mei Zhang, Shaojun Xie, Jian Chen, Weibin Song, Xin Wang and Silong Sun for helpful discussions concerning the manuscript. This work was supported by grants from the 863 Project (2012AA10A305) and the National Natural Science Foundation of China (31421005, 31225020 and 91435206).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.pbi.2015.01.008>.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Abbo S, Pinhasi Van-Oss R, Gopher A, Saranga Y, Ofner I, Peleg Z: **Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes.** *Trends Plant Sci* 2014, **19**:351-360.
2. Meyer RS, Purugganan MD: **Evolution of crop species: genetics of domestication and diversification.** *Nat Rev Genet* 2013, **14**:840-852.
3. Gross BL, Olsen KM: **Genetic perspectives on crop domestication.** *Trends Plant Sci* 2010, **15**:529-537.
4. Morrell PL, Buckler ES, Ross-Ibarra J: **Crop genomics: advances and applications.** *Nat Rev Genet* 2012, **13**:85-96.
5. Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD, Gaut BS: **The effects of artificial selection on the maize genome.** *Science* 2005, **308**:1310-1314.
6. Doebley JF, Gaut BS, Smith BD: **The molecular genetics of crop domestication.** *Cell* 2006, **127**:1309-1321.
7. Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L *et al.*: **Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes.** *Nat Biotechnol* 2011, **30**:105-111.
8. Lai J, Li R, Xu X, Jin W, Xu M, Zhao H, Xiang Z, Song W, Ying K, Zhang M *et al.*: **Genome-wide patterns of genetic variation among elite maize inbred lines.** *Nat Genet* 2010, **42**:1027-1030.
9. Nielsen R: **Molecular signatures of natural selection.** *Annu Rev Genet* 2005, **39**:197.
10. Nielsen R: **Genomic scans for selective sweeps using SNP data.** *Genome Res* 2005, **15**:1566-1575.
11. Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ *et al.*: **Detecting recent positive selection in the human genome from haplotype structure.** *Nature* 2002, **419**:832-837.
12. Huang X, Kurata N, Wei X, Wang Z, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W *et al.*: **A map of rice genome variation reveals the origin of cultivated rice.** *Nature* 2012, **490**:497-501.
- This paper described the patterns of genetic changes during rice domestication and proposed a single domestication hypothesis that *O. sativa* ssp. *japonica* was first domesticated, and then *O. sativa* ssp. *indica* was generated with crosses between *japonica* and local wild accessions.
13. Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y, Gu X *et al.*: **A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity.** *Nat Genet* 2013, **45**:1510-1515.
14. Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, Zhang Z, Lun Y, Li S, Wang X *et al.*: **Genomic analyses provide insights into the history of tomato breeding.** *Nat Genet* 2014, **46**:1220-1226.
15. Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L: **Natural selection has driven population differentiation in modern humans.** *Nat Genet* 2008, **40**:340-345.
16. Chen H, Patterson N, Reich D: **Population differentiation as a test for selective sweeps.** *Genome Res* 2010, **20**:393-402.
17. Ross-Ibarra J, Morrell PL, Gaut BS: **Plant domestication, a unique opportunity to identify the genetic basis of adaptation.** *Proc Natl Acad Sci U S A* 2007, **104**(Suppl. 1):8641-8648.
18. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia J, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaepler SM *et al.*: **Comparative population genomics of maize domestication and improvement.** *Nat Genet* 2012, **44**:808-811.
- In this paper, the authors identified the genomic signals during maize domestication and breeding and revealed that selection during maize domestication can happen in regions without genes.
19. Gepts P: **The contribution of genetic and genomic approaches to plant domestication studies.** *Curr Opin Plant Biol* 2014, **18**:51-59.
20. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D *et al.*: **Genetic structure and domestication history of the grape.** *Proc Natl Acad Sci U S A* 2011, **108**:3530-3535.
21. Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, Zheng Y, Mao L, Ren Y, Wang Z *et al.*: **The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions.** *Nat Genet* 2012, **45**:51-58.
22. Harushima Y, Nakagahra M, Yano M, Sasaki T, Kurata N: **Diverse variation of reproductive barriers in three intraspecific rice crosses.** *Genetics* 2002, **160**:313-322.
23. Lam H, Xu X, Liu X, Chen W, Yang G, Wong F, Li M, He W, Qin N, Wang B *et al.*: **Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection.** *Nat Genet* 2010, **42**:1053-1059.
- The authors observed an excess of low-frequency alleles in cultivated soybean, which indicates population expansion after domestication.

24. Jiao Y, Zhao H, Ren L, Song W, Zeng B, Guo J, Wang B, Liu Z, • Chen J, Li W *et al.*: **Genome-wide genetic changes during modern breeding of maize**. *Nat Genet* 2012, **44**:812-815.
The authors reported the types and rates of genetic changes during modern maize breeding, identified artificial selective sweeps in both genic and non-genic regions and observed the increasing proportion of rare alleles with the maize breeding process.
25. Li C, Zhou A, Sang T: **Rice domestication by reducing shattering**. *Science* 2006, **311**:1936-1939.
26. Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Fallier M, Bombliès K, Lukens L, Doebley JF: **The origin of the naked grains of maize**. *Nature* 2005, **436**:714-719.
27. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD: **fw2.2: a quantitative trait locus key to the evolution of tomato fruit size**. *Science* 2000, **289**: 85-88.
28. Tian Z, Wang X, Lee R, Li Y, Specht JE, Nelson RL, McClean PE, Qiu L, Ma J: **Artificial selection for determinate growth habit in soybean**. *Proc Natl Acad Sci U S A* 2010, **107**:8563-8568.
29. Salvi S, Sponza G, Morgante M, Tomez D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashv S, Bruggemann E *et al.*: **Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize**. *Proc Natl Acad Sci U S A* 2007, **104**:11376-11381.
30. Studer A, Zhao Q, Ross-Ibarra J, Doebley J: **Identification of a functional transposon insertion in the maize domestication gene *tb1***. *Nat Genet* 2011, **43**:1160-1163.
31. Zhou L, Zhang J, Yan J, Song R: **Two transposable element insertions are causative mutations for the major domestication gene *teosinte branched 1* in modern maize**. *Cell Res* 2011, **21**:1267-1270.
32. Li X, Zhu C, Yeh CT, Wu W, Takacs EM, Petsch KA, Tian F, Bai G, • Buckler ES, Muehlbauer GJ *et al.*: **Genic and nongenic contributions to natural variation of quantitative traits in maize**. *Genome Res* 2012, **22**:2436-2444.
A report of the systematic tabulation of the relative positions of GWAS trait-associated SNPs (TASs) in maize. About half of the TASs were reported to locate in the 5'-upstream promoter regions of genes.
33. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J *et al.*: **Systematic localization of common disease-associated variation in regulatory DNA**. *Science* 2012, **337**:1190-1195.
34. Wang M, Yu Y, Haberer G, Marri PR, Fan C, Goicoechea JL, • Zuccolo A, Song X, Kudrna D, Ammiraju JSS *et al.*: **The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication**. *Nat Genet* 2014, **46**:982-988.
The authors assembled the *O. glaberrima* genome and gave an example of convergent domestication between *O. glaberrima* and *O. sativa*.
35. Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J: **A single domestication for maize shown by multilocus microsatellite genotyping**. *Proc Natl Acad Sci U S A* 2002, **99**:6080-6084.
36. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu S, Song Q, Chavarro C *et al.*: **A reference genome for common bean and genome-wide analysis of dual domestications**. *Nat Genet* 2014, **46**:707-713.
37. Mace ES, Tai S, Gilding EK, Li Y, Prentis PJ, Bian L, Campbell BC, Hu W, Innes DJ, Han X *et al.*: **Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum**. *Nat Commun* 2013, **4**:2320.
38. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A *et al.*: **Finding the missing heritability of complex diseases**. *Nature* 2009, **461**:747-753.
39. Springer NM: **Epigenetics and crop improvement**. *Trends Genet* 2013, **29**:241-247.
40. Zhang X: **The epigenetic landscape of plants**. *Science* 2008, **320**:489-492.
41. Law JA, Jacobsen SE: **Establishing, maintaining and modifying DNA methylation patterns in plants and animals**. *Nat Rev Genet* 2010, **11**:204-220.
42. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urlich MA, Libiger O, Schork NJ, Ecker JR: **Transgenerational epigenetic instability is a source of novel methylation variants**. *Science* 2011, **334**:369-373.
43. Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M: **The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana***. *Science* 2009, **327**:92-94.
44. Zhang L, Cheng Z, Qin R, Qiu Y, Wang JL, Cui X, Gu L, Zhang X, Guo X, Wang D *et al.*: **Identification and characterization of an epi-allele of *FIE1* reveals a regulatory linkage between two epigenetic marks in rice**. *Plant Cell* 2012, **24**:4407-4421.
45. Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB: **A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening**. *Nat Genet* 2006, **38**:948-952.
46. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D: **Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome**. *Nature* 2011, **480**:245-249.
47. Eichten SR, Briskine R, Song J, Li Q, Swanson-Wagner R, • Hermanson PJ, Waters AJ, Starr E, West PT, Tiffin P *et al.*: **Epigenetic and genetic influences on DNA methylation variation in maize populations**. *Plant Cell* 2013, **25**:2783-2797.
The authors reported the DNA methylome analysis of 20 founder inbred lines of a maize NAM population and concluded that DNA methylation variation is influenced by genetic and epigenetic changes that can often be stably inherited in offspring.
48. Schmitz RJ, Schultz MD, Urlich MA, Nery JR, Pelizzola M, • Libiger O, Alix A, McCosh RB, Chen H, Schork NJ *et al.*: **Patterns of population epigenomic diversity**. *Nature* 2013, **495**:193-198.
Through a population-scale whole genome MethylC-seq in *Arabidopsis*, the authors reported naturally existing DNA-methylation variants, which could be the target of future crop breeding.
49. Schmitz RJ, He Y, Valdes-Lopez O, Khan SM, Joshi T, Urlich MA, Nery JR, Diers B, Xu D, Stacey G *et al.*: **Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population**. *Genome Res* 2013, **23**:1663-1674.
50. Chodavarapu RK, Feng S, Ding B, Simon SA, Lopez D, Jia Y, Wang GL, Meyers BC, Jacobsen SE, Pellegrini M: **Transcriptome and methylome interactions in rice hybrids**. *Proc Natl Acad Sci U S A* 2012, **109**:12040-12045.
51. Shen H, He H, Li J, Chen W, Wang X, Guo L, Peng Z, He G, Zhong S, Qi Y *et al.*: **Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids**. *Plant Cell* 2012, **24**:875-892.
52. Greaves IK, Groszmann M, Ying H, Taylor JM, Peacock WJ, Dennis ES: **Trans chromosomal methylation in *Arabidopsis* hybrids**. *Proc Natl Acad Sci U S A* 2012, **109**:3570-3575.
53. Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, • Labadie K, Caillieux E, Aury J, Wincker P, Roudier F: **Mapping the epigenetic basis of complex traits**. *Science* 2014, **343**: 1145-1148.
This paper demonstrated that six DMRs act as epi-QTLs, which account for a large proportion (60–90%) of heritability of two complex traits in *Arabidopsis*, which clearly demonstrates the relationship between epigenetic modification and complex traits in plants.
54. Rakyán VK, Down TA, Balding DJ, Beck S: **Epigenome-wide association studies for common human diseases**. *Nat Rev Genet* 2011, **12**:529-541.
An excellent review introducing the EWAS design, sample selection and the integration of the GWAS and EWAS results for dissecting complex GWAS signals.
55. Breittling LP, Yang R, Korn B, Burwinkel B, Brenner H: **Tobacco-smoking-related differential DNA methylation: 27 K discovery and replication**. *Am J Hum Genet* 2011, **88**:450-457.
56. Bell JT, Tsai P, Yang T, Pidsley R, Nisbet J, Glass D, Mangino M, Zhai G, Zhang F, Valdes A *et al.*: **Epigenome-wide scans identify differentially methylated regions for age and age-related**

- phenotypes in a healthy ageing population.** *PLoS Genet* 2012, **8**:e1002629.
57. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, Meduri E, Morange P, Gagnon F, Grallert H *et al.*: **DNA methylation and body-mass index: a genome-wide analysis.** *Lancet* 2014, **383**:1990-1998.
 58. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C *et al.*: **Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci.** *Nat Neurosci* 2014, **17**:1156-1163.
 59. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA *et al.*: **The B73 maize genome: complexity, diversity, and dynamics.** *Science* 2009, **326**:1112-1115.
 60. Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D *et al.*: **Control of a key transition from prostrate to erect growth in rice domestication.** *Nat Genet* 2008, **40**:1360-1364.
 61. Jin J, Huang W, Gao J, Yang J, Shi M, Zhu M, Luo D, Lin H: **Genetic control of rice plant architecture under domestication.** *Nat Genet* 2008, **40**:1365-1369.
 62. Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H *et al.*: **Control of rice grain-filling and yield by a gene with a potential signature of domestication.** *Nat Genet* 2008, **40**:1370-1374.
 63. Doebley J, Stec A, Gustus C: **Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance.** *Genetics* 1995, **141**:333-346.
 64. Lin Z, Li X, Shannon LM, Yeh C, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE *et al.*: **Parallel domestication of the Shattering1 genes in cereals.** *Nat Genet* 2012, **44**:720-724.
 65. Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q: **GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein.** *Theor Appl Genet* 2006, **112**:1164-1171.
 66. Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J *et al.*: **Natural variation in GS5 plays an important role in regulating grain size and yield in rice.** *Nat Genet* 2011, **43**:1266-1269.
 67. Song X, Huang W, Shi M, Zhu M, Lin H: **A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase.** *Nat Genet* 2007, **39**:623-630.
 68. Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X *et al.*: **Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight.** *Cell Res* 2008, **18**:1199-1209.
 69. Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, Xia G, Chu C, Li J, Fu X: **Natural variation at the DEP1 locus enhances grain yield in rice.** *Nat Genet* 2009, **41**:494-497.
 70. Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X *et al.*: **Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice.** *Nat Genet* 2008, **40**:761-767.
 71. Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F *et al.*: **Control of tillering in rice.** *Nature* 2003, **422**:618-621.
 72. Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X *et al.*: **Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice.** *Nat Genet* 2010, **42**:541-544.